SELECTION OF INTELLIGENCE ANALYSTS

M. A. Fischl

and

Arthur C. F. Gilbert

U.S. Army Research Institute for the
Behavioral and Social Sciences

The Defense Intelligence Agency (DIA) is responsible for keeping the Joint Chiefs of Staff informed of activities of potential military consequence world-wide. A very important element in performance of this mission is the uniformed and civilian staff of individuals who analyze social, political, economic, and strategic information about assigned parts of the world.

Civilian intelligence analysts are selected for their jobs largely on the basis of academic record and prior work history. Virtually all are college graduates, many with graduate training and degrees, and many are former uniformed military personnel. This paper describes a recent investigation into the feasibility of improving the process of selection of civilian intelligence analysts, through adding the use of tests of the aptitudes and skills required in the job.

Method

The method employed followed a standard test-development paradigm. Job analysis identified personal characteristics important to analyst success, an experimental battery of tests to measure the characteristics was selected, and it was administered to a sample of recently-hired incumbents for whom job performance information was also obtained. Multiple regression analyses weighted the tests, which were then cross-validated on holdout portions of the sample.

Job Analysis

Discussions were held with members of the DIA staff, to learn the nature of the job performed by intelligence analysts and apparent causes of success and failure on the job. Additional information about the analyst job and characteristics judged important for its successful performance was obtained from personal interviews with 14 incumbent intelligence analysts and from critical incident questionnaires on which 20 supervisors of analysts provided descriptions of positive and negative critical incidents, with explanation of the personal qualities responsible for each incident.

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

Experimental Test Battery

On the basis of the job analysis, a picture emerged of personal attributes important in the intelligence analyst job. Table 1 presents these attributes and the commercial tests selected to measure them and serve as potential predictors of analyst success.

Table 1
Experimental Predictor Variables
and Their Tests

Variable	Test
High Level reasoning ability	Watson-Glaser Critical Thinking Appraisal 1/
Inductive reasoning	Comprehensive Ability Battery: Subtest 6, Inductive Reasoning 2/
Intellectual flexibility	Comprehensive Ability Battery: Subtest 15, Spontaneous Flexibility
Writing skill	Flanagan Industrial Tests: Subtest 6, Expression 3/
Memory	Flanagan Industrial Tests: Subtest 12, Memory
Intellectual curiosity	Gordon Personal Inventory: Original Thinking Scale 1/
Deliberateness, carefulness	Gordon Personal Inventory: Cautiousness Scale
Interpersonal skill	Gordon Personal Inventory: Personal Relations Scale
Achievement motivation	Gordon Survey of Personal Values: Achievement Scale 3/
Self-discipline	Gordon Survey of Personal Values: Orderliness Scale
Perserverance	Gordon Personal Profile: Responsiblity Scale 1/

1/ New York: The Psychological Corporation

2/ Champaign, IL: Institute for Personality and Ability Testing

3/ Chicago: Science Research Associates

Subjects and Procedure

The experimental battery was administered in a 3-hour session to 64 intelligence analysts who had been employed at DIA for periods ranging from 1 to 24 months. The mean experience level was just under 12 months, and the sample was approximately 2/3 male, 1/3 female. All but 3 members of the sample were caucasian. These 64 analysts were the most recently hired by DIA.

Immediately after the testing session a DIA staff member met individually with the supervisors of the 64 analysts to administer performance rating forms. At that time the supervisors were informed that the ratings were *ad hoc* and for research purposes only, would not appear on a personnel record, and, when completed, were to be transmitted in sealed envelopes directly to the research organization outside of DIA. Candor and accuracy were encouraged, it being pointed out that there was no risk to any employee but potential great benefit to the Agency. A copy of the rating form appears as Figure 1.

ANALYST RATING FORM

Date: _____

Analyst's Name (Print)

_____
Last First MI

Rater's Name (Print)

_____
Last First MI

Length of time you have known this analyst _____ (Months)

INSTRUCTIONS

COMPARED TO ALL ANALYSTS YOU HAVE KNOWN, using the rating scale below, blacken one box to indicate your appraisal of this analyst's performance. In making your rating of performance, consider the analyst's demonstrated performance relative to the performance of all other analysts you have known at his/her stage of experience.

Bottom 10% (Marginal)	Next higher 20%	Middle 40% (Average)	Next higher 20%	Top 10% (Outstanding)
1	2	3	4	5

Figure 1. Performance appraisal instrument

The full sample of 64 intelligence analysts was randomly divided into two half samples, and independent stepwise multiple regression analyses were performed on each. The regression weights emerging from analysis of half-sample A were utilized to compute a score for each member of half-sample B, and that score distribution was correlated with the distribution of criterion ratings for these individuals. Similarly, the regression weights emerging from analysis of half-sample B were utilized to compute a score for each member of half-sample A, and that score distribution was correlated with the distribution of their criterion ratings.

The validation procedure will be recognized as standard double cross-validation, yielding two regression equations and two validity coefficients. From that point, judgment was utilized to integrate the two solutions-- that is, to select the final test battery-- and to arrive at a single best estimate of the criterion-related validity of that battery.

Results

The most valid test was the test of Expression, correlating 0.55 with the criterion in half-sample A and 0.37 in half-sample B. Of equal validity to Expression in half-sample B was the Critical Thinking Appraisal, and this test also correlated 0.36 with the criterion in half-sample A.

Addition of tests resulted in five-variable solutions in both half-samples. In half-sample A this solution was:

$$Y = 2.144 + 0.054 \text{ Expression} - 0.036 \text{ Orderliness} + 0.015 \text{ Memory} + 0.015 \text{ Spontaneous Flexibility} + 0.010 \text{ Critical Thinking}$$

In half-sample B the five-variable solution was:

$$Y = 1.36 + 0.046 \text{ Expression} + 0.026 \text{ Critical Thinking} - 0.025 \text{ Orderliness} + 0.022 \text{ Memory} + 0.013 \text{ Spontaneous Flexibility}$$

When these five-variable regression equations were cross-validated in the opposite half-samples, the resulting correlation coefficients were 0.60 and 0.38.

Note was taken that Orderliness had a negative regression weight in both solutions. For operational application the use of negative weights was judged highly undesirable, and the Orderliness scale was deleted from further consideration. At the same time, more careful examination of the Spontaneous Flexibility test disclosed that it could not be scored by a non-professional, that careful subjective judgment was needed. This second administrative concern disqualified the test of Spontaneous Flexibility. The remaining three variables were common to both regression equations, and the only tasks remaining were to derive a new set of weights for a three-variable equation, and to estimate the validity of the three-test battery. These tasks were performed in each half-sample, crossed on the other, and also in the full sample of 64 cases. Table 2 presents the outcomes.

Table 2
Regression Weights and Validity Coefficients
for the Three-Test Battery

Weights	Validity Coefficient
0.069 Expression + 0.011 Critical Thinking + 0.009 Memory	0.42, N_A= 32
0.045 Expression + 0.032 Critical Thinking + 0.003 Memory	0.54, N_B =32
0.057 Expression + 0.021 Critical Thinking + 0.007 Memory	0.50, N =64

Discussion

The three-test battery consisting of the Watson-Glaser Critical Thinking Appraisal, and the Memory and Expression subtests from the series of Flanagan Industrial Tests, requires about 1½-hours for administration and is scorable by a clerk using stencil overlays.

If simple unit weights--i.e., the sum of raw scores on the three tests--are employed in operational use of the battery, the counterparts to the validity coefficients shown in Table 2 become respectively 0.41, 0.48, and 0.44, all significant at $p \leq 0.01$.

Analysis to detect any adverse gender impact disclosed no difference in test battery scores of the women and the men in the sample. Using the simpler weights, for which the total possible score is 150, the women's mean was 105 and the men's mean was 104.

An alternative to the unit weighting procedure might be differential whole-number weighting of the tests of the battery. Inasmuch as the regression analyses weighted Expression between 1½ and 7 times as heavily as Critical Thinking and between 7 and 15 times as heavily as Memory, a set of weights in the ratio 5:2:1 or 6:2:1 might be superior. Its use was not investigated.

On the basis of the investigation performed, it appears that a relatively short battery of easily administered and scored tests can appreciably improve the procedure for selection of intelligence analysts.

Acknowledgment

The authors acknowledge the assistance of Ms. Donna Angle in this activity.